Genetics, Disease and Race

Timothy R. Rebbeck

Center for Clinical Epidemiology and Biostatistics
University of Pennsylvania School of Medicine

CCEB

Is "Race" a Biologically Meaningful Entity?

- **"There is no such thing as a biological entity that warrants the term 'race'."**
  - Dr. C. Loring Brace

- **"Forensic anthropologists are overwhelmingly in support of the idea of the basic biological reality of human races…"**
  - Dr. George Gill

(Source: NOVA Special "Does Race Exist?")

Premise

- "Race" is widely used in research.
- There is more genetic variability within "races" than between them.
- "Race" remains a poor index for many entities of interest: better measures should be developed.
- Interpretation of research results should distinguish data involving race from value judgments about race.

When Might It Be Appropriate to Consider Race in Genetics Research?

1. Targeted Gene Discovery
2. Study Design
3. Bias and Confounding
4. Genome Scans/Mapping

CCEB

Targeted Gene Discovery

- **Race may be a surrogate for identifying genetically high risk groups (cf. family history)**
- **Race may identify exposed-predisposed groups**
- **Founder mutations may exist in genetically homogeneous groups (e.g., French Canadians, Icelanders, Ashkenazi Jews, etc.)**

When Might It Be Appropriate to Consider Race in Genetics Research?

1. Targeted Gene Discovery
2. Study Design
3. Bias and Confounding
4. Genome Scans/Mapping

CCEB

Study Design: Genotype Frequencies Differ by Race

Group	*CYP3A4*1B Frequency*
Asian	0%
Caucasian	7%
Latino	11%
African American	58%
African	77%

***Problem*: If genotype frequency estimates ignore ethnicity, study may have inadequate power or be inefficient.**

When Might It Be Appropriate to Consider Race in Genetics Research?

1. Targeted Gene Discovery
2. Study Design
3. Bias and Confounding
4. Genome Scans/Mapping

CCEB

Bias and Confounding

IF: Baseline disease risks differ across ethnicities

AND: Risk-conferring allele frequencies differ across ethnicities

THEN: Confounding by ethnicity (population stratification) may result that can produce biased effect estimates

CCEB

CYP3A4, Race, and Prostate Cancer Risk

Group	Prostate Cancer Rate per 100,000	*CYP3A4*1B* Frequency
Asian	2	0%
Caucasian	101	7%
African American	137	58%

The Effects of Population Stratification are Limited

- Small if baseline disease or allele frequency differences between races are small
- Diminishes as admixture increases
- May be more pronounced in recently admixed populations
- Adjustment can be undertaken when “race” can be measured
- Poor study design may be a greater problem

When Might It Be Appropriate to Consider Race in Genetics Research?

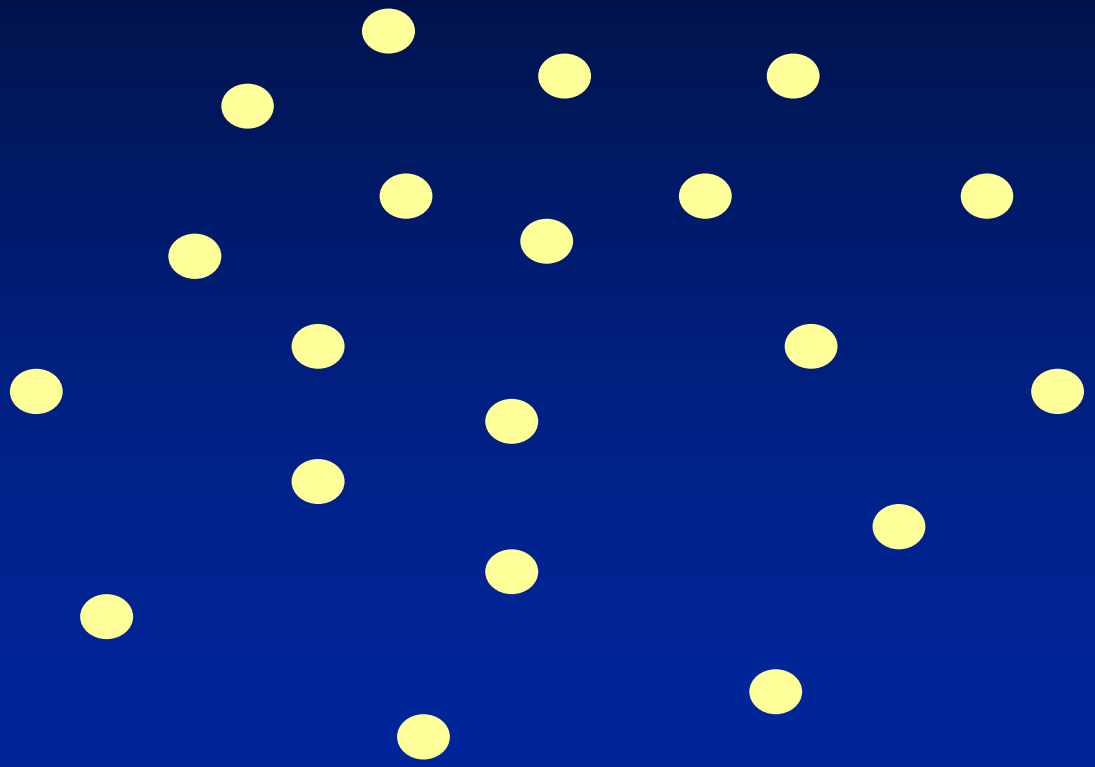
1. Targeted Gene Discovery
2. Study Design
3. Bias and Confounding
4. Genome Scans/Mapping

CCEB

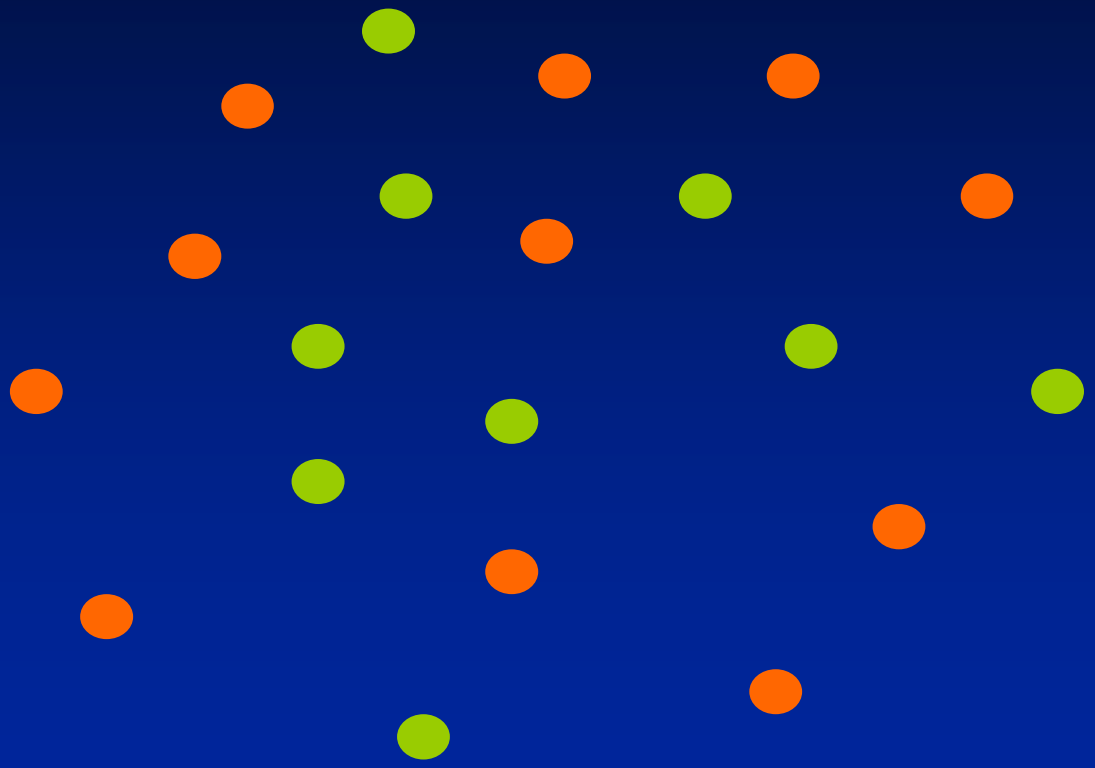
Genome Scans, Gene Mapping

1. **Methods rely on relationships among genetic variants on a chromosome.**
2. **Differences in the pattern or frequency of genetic variants will affect research results.**
3. **The pattern and frequency of genetic variants varies by race.**

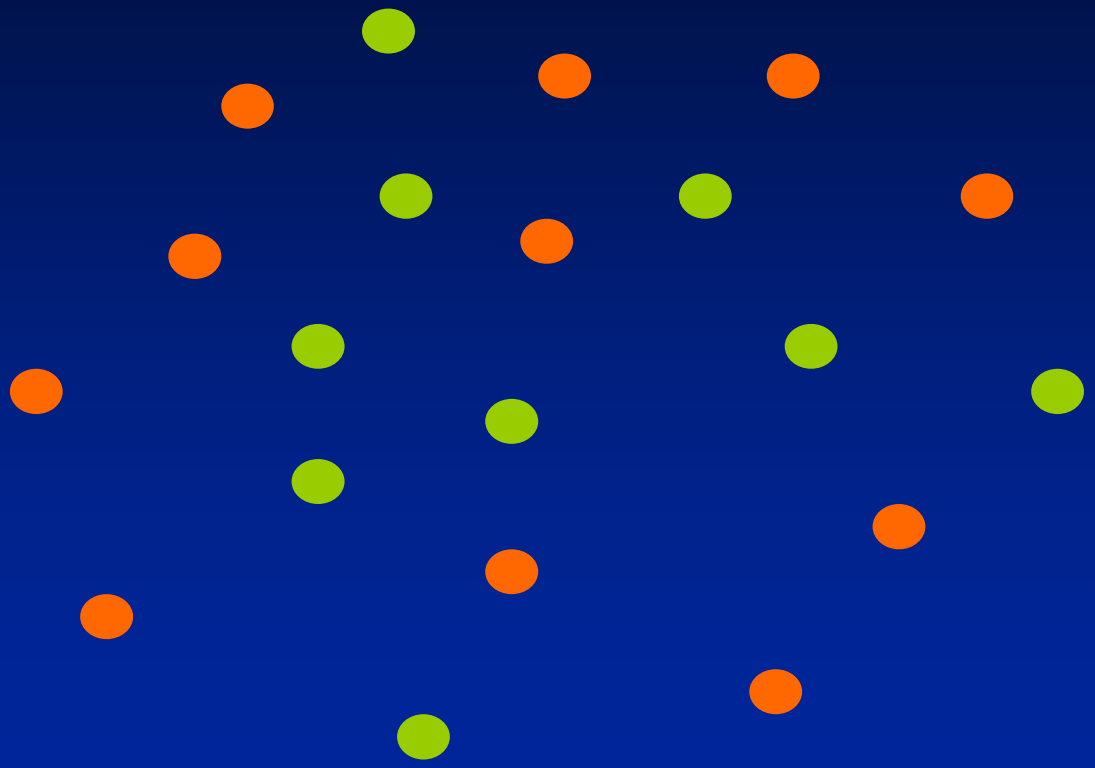
Assemble Sample Set for Population Genetics Analysis



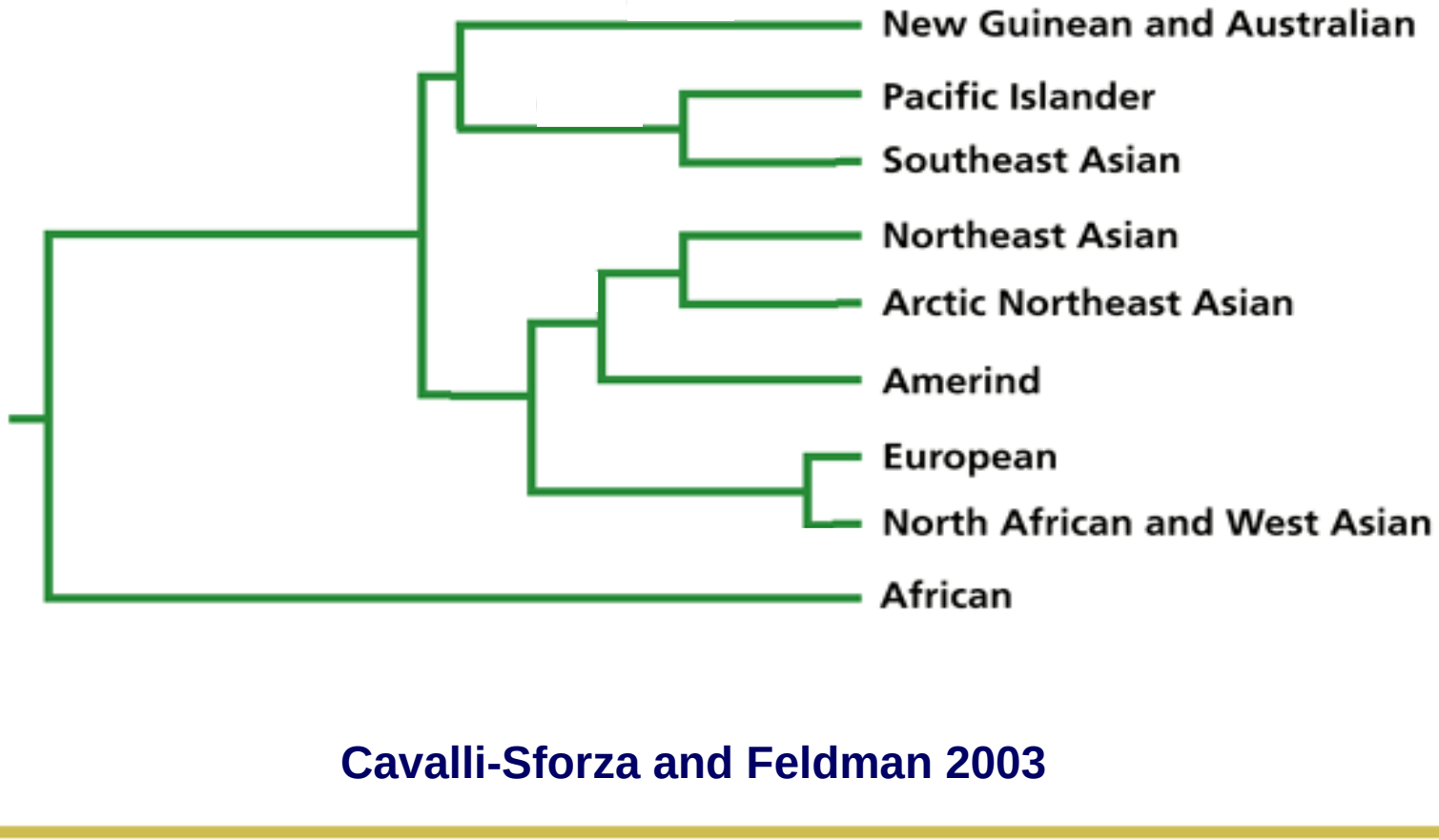
Genotype



Cluster by Genetic Similarity

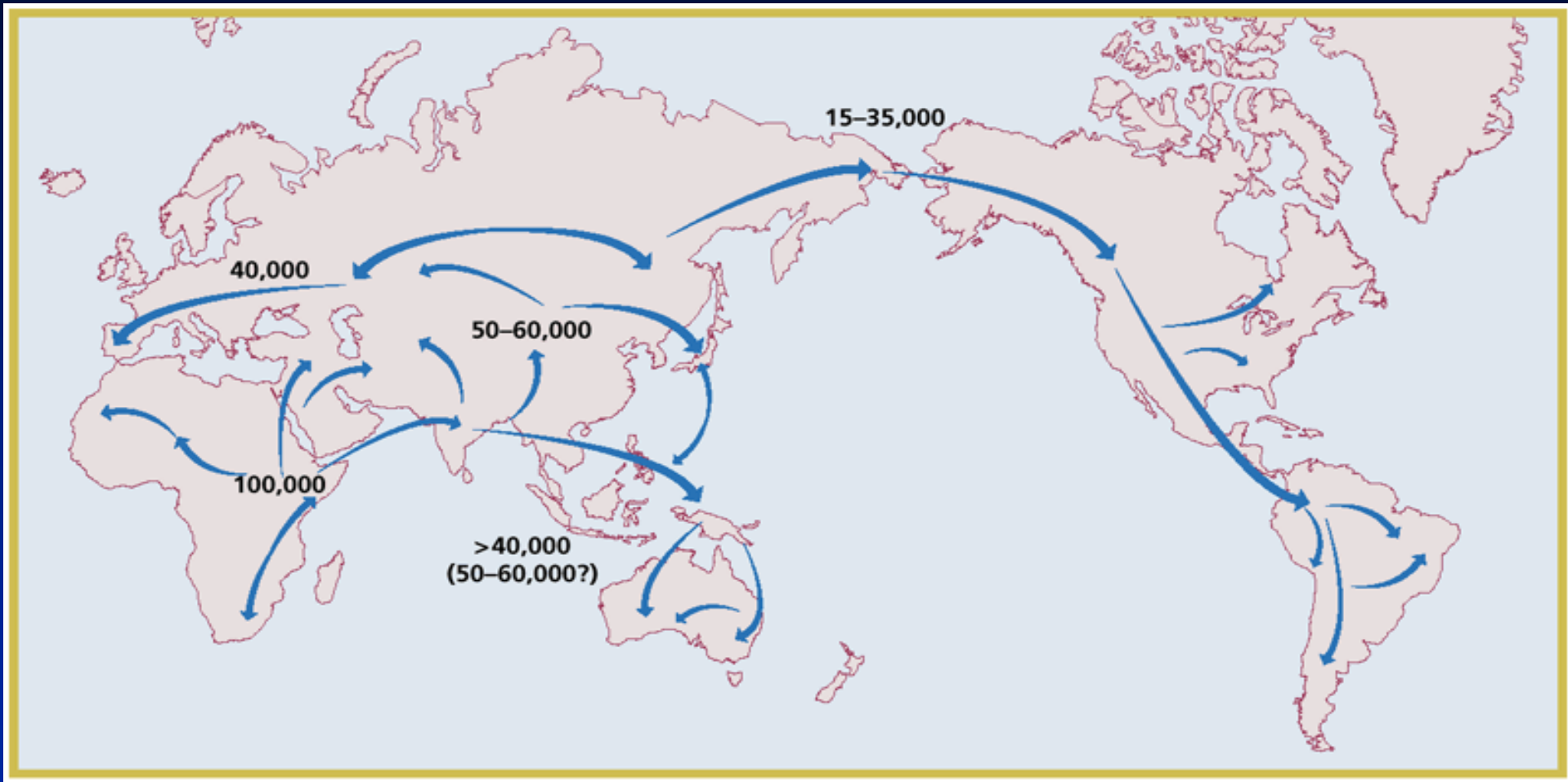


Genetic Diversity and Race



Cavalli-Sforza and Feldman 2003

Human Migration History



Cavalli-Sforza and Feldman 2003

CYP3A Haplotypes:
CYP3A4-CYP3A5-CYP3A43

Haplotype	Frequency in: CA	AF
1.1.1	.838	.247
1.1.2	.058	.025
1.2.1	.038	.064
1.2.2	.003	.044
2.1.1	.030	.098
2.1.2	.002	.045
2.2.1	.030	.278
2.2.2	.002	.198

Genomic Differences and Race

- Racial differences in genome structure are consistent and correlated with:
  - Geography
  - Age and origin of humans
  - Cultural data such as language
- Greater genetic diversity within populations than between populations

Summary

- Information about “race” can aid research.
- Ignoring “race” can lead to biased results or inefficient approaches.
- “Race” remains a poor index for many entities of interest: better measures should be developed.
- Interpretation of research results should distinguish data involving race from value judgments about race.